# Developement of Bayesian Networks from Unified Modeling Language for Learner Modelling

ANOUAR TADLAOUI Mouenis
LIROSA
Faculty of Sciences, Abdelmalek Essaadi University
Tétouan, Morocco

AAMMOU Souhaib
LIROSA
Faculty of Sciences, Abdelmalek Essaadi University
Tétouan, Morocco

KHALDI Mohamed
LIROSA
Faculty of Sciences, Abdelmalek Essaadi University
Tétouan, Morocco

***Abstract*—First of all, and to clarify our purpose, it seems important to say that the work we are presenting here lie within the framework of learner modeling in an adaptive system understood as computational modeling of the learner .we must state also that Bayesian Networks are effective tools for learner modeling under uncertainty. They have been successfully used in many systems, with different objectives, from the assessment of knowledge of the learner to the recognition of the plan followed in problem solving. The main objective of this paper is to develop a Bayesian networks for modeling the learner from the use case diagram of the Unified Modeling Language. To achieve this objective it is necessary first to ask the Why and how we can represent a Learner model using Bayesian networks? How can we go from a dynamic representation of the learner model using UML to a probabilistic representation with Bayesian networks? Is this approach considered experimentally justified? First, we will return to the definitions of the main relationships in the diagram use cases and Bayesian networks, and then we will focus on the development rules on which we have based our work. We then demonstrate how to develop a Bayesian network based on these rules. Finally we will present the formal structure for this consideration. The prototypes and diagrams presented in this work are arguments in favor of our objective. And the network obtained also promotes reusing the learner modeling through similar systems.**

***Keywords—Learner Modeling; Bayesian networks; Cognitive diagnosis; Uncertainty***

## I. Introduction

The problem of this paper can be summarized as follows: Why and how can we represent a Learner model using Bayesian networks? How can we go from a dynamic representation of the Unified modeling language (UML) model to a probabilistic representation with Bayesian networks? Is this consideration experimentally justified?

The learner model is a data structure that represents the state of knowledge of a learner in a given field. This model identifies the learner's current level of understanding of the domain knowledge. It includes data on individual variables of a learner that allow updating of the learner profiles from information obtained during the interactions.

All existing approaches to model the learner are based generally on using the Unified Modeling Language [1], that quickly became a standard for the analysis and design in software development. It provides a schematic approach to describing the needs of the user, which begins with the use cases diagrams, and leads to a more formal specification, using stereotyped classes in the analysis model. The components of this modeling language form the basis of an architectural view in the system while providing the foundation for the design, implementation and validation and verification.

We have attempted in previous works, to model the learner using Bayesian networks [2] and multi networks [3] as a formalism to manage uncertainty in the management of learner model. In this paper, we will try to offer a combination of these two approaches, starting with specifying the transformation rules on which we have based our work. We will then demonstrate how to transform the use case diagrams into a Bayesian network based on these rules. Finally we will present the formal structure for this consideration.

## II. Use Case Diagram, UML Point of View

### A. Definition

Use cases describe the form of actions, reactions and the behavior of a system from a user perspective. They allow defining the limits of the system and the relationship between the system and the environment.

Use cases are filling a lack of raw object methods, such as [4] and [5], which did not offer techniques, for the identification of needs. In this sense, the use cases associated with technical objects allow a comprehensive approach to the entire life cycle, from the specification to implementation.

A use case is a specific way of using a system. It is the image of system functionality, triggered in response to the stimulation of an external actor.

### B. Use case diagram's Relationships

Use case diagrams represent actors and relationships between actions and actors. We will define in this section the main relationships in the use case diagram that we will use in our work, they are three types of relationships between actors and use cases:

#### 1) Generalization relationship

A case A is a generalization of a case B if B is a particular case of A. For example, the consultation of an account via the Internet is a particular case of the consultation. This relationship of generalization / specialization is present in most of the UML diagrams and results in the concept of inheritance in object-oriented languages.

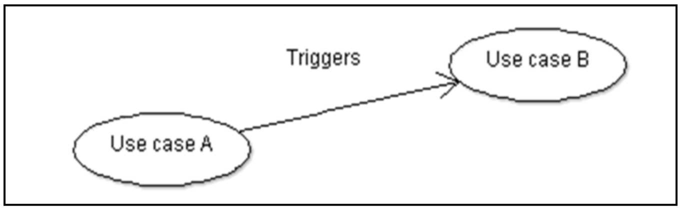

Fig. 1. Generalization relationship in a use case diagram

“Figure. 1”, shows that use case A initiates a use case B; this action is represented by an arrow from the initiator of the action to the triggered action.

#### *2) Inclusion relationship*

A case A includes a case B if the behavior described by the case A includes the conduct of the case B: Where A depends on B. When A is applied, B also must be applied as a part of A.

This dependence is symbolized by the "include" stereotype. For example, accessing information from a bank account necessarily includes an authentication phase with a username and password.

The inclusions essentially allow factorizing a part of the description of a use case that would be common to other use cases like "Figure.2" shows. The inclusions are also used to decompose a complex case into simpler sub-cases.

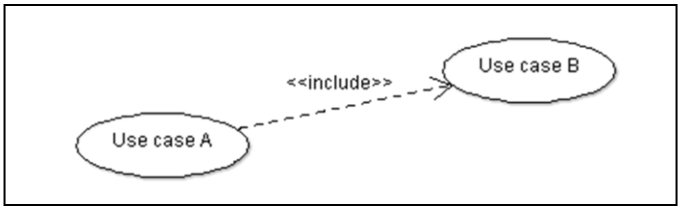

Fig. 2. Inclusion relationship in a use case diagram

#### *3) Extension relationship*

It is said that use case A extends a use case B when the use case A can be called during the execution of the use case B. Running B may possibly lead to the execution of A: unlike the inclusion, the extension is optional. This dependence is symbolized by the stereotype "extend"

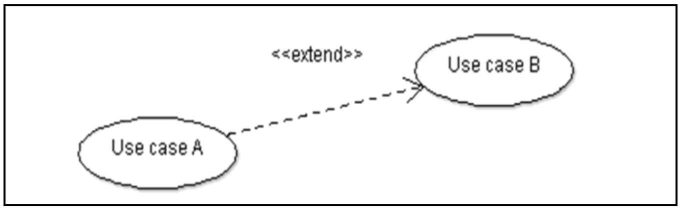

Fig. 3. Extension relationship in a use case diagram

## III. LEARNER MODEL

### *A. Definition*

A learner model allows keeping the learner information, for example his level of knowledge on a given topic (performance), his frequent mistakes/misunderstandings, psychological characteristics, etc.

A learner model can be defined as a set of structured information about the learning process, and this structure contains values on the characteristics of the learner. [6] It provides the necessary data to the other modules to achieve the adaptation of teaching to the learner. [7]

Many studies emphasize the uncertainty of the information contained in the student model and the importance of the intention behind the creation of this model. Thus, a student model represents the belief system about learners' beliefs (the system's beliefs about the learner's beliefs) accumulated during the diagnostic process.

### *B. Typologies of Learner model*

The learner model is a data structure in the computer sense that characterizes for the learning environment, the state of a subset of the learner’s knowledge from the system point of view.

It will be defined by the difference between the learner knowledge and target knowledge, issue of learning, as represented in the system. The approach to represent this difference leads to distinguish two major classes of models:

- The models of partial or overlay expertise [8], in which the knowledge of the learner is only a subset of the target knowledge. The idea behind this type of model is that the learner present deficiencies or the poorly insured knowledge, or somehow weaknesses, it is identified to allow it to grow. The aim of the learning system is then to complete the knowledge of the learner in order to acquire all the knowledge outlined in the model.
- Differential models [9], which incorporate "false knowledge", corresponding to perturbations of the expert knowledge or erroneous preconceptions. In fact, studies show that many errors are not due to erratic behavior of learners, but the correct application of false procedures. To develop a model of learners' knowledge, one must take into account these types of systematic errors, that researchers will be designated by the term "bug" (bug).

While a partial model invites expertise in teaching strategies centered on the fact to fill the gaps of the learner, the incremental models will lead to strategies based on remediation.

## IV. BAYESIAN NETWORKS

### *A. Definition*

Knowledge representation and reasoning from these representations has created many models. Probabilistic graphical models, specifically Bayesian networks initiated by [10] in the 1980s, have proven to be useful tools for representing uncertain knowledge and reasoning from incomplete information.

A Bayesian network $B = (G, N)$ is defined by

- $G = (X, E)$ , acyclic directed graph with vertices associated with a set of random variables $X = (X, \dots, Xn)$ ;

- $N = \{P(Xi \mid Pa(Xi))\}$ All the probabilities of each node $Xi$ conditionally to the state of its parents $Pa(Xi)$ in G.

Thus, the graphical part of the Bayesian network indicates the dependencies (or independence) between variables and provides a visual tool for knowledge representation, more easily comprehensible by its users. In addition, the use of probability allows taking into account uncertainty in quantifying the dependencies between variables. Both properties have been the cause of the first names of Bayesian networks, "probabilistic expert systems", where the graph was compared to the set of rules of conventional expert system, and the conditional probabilities presented as a quantification uncertainty about the rules.

[11] Also have shown that Bayesian networks allow representing compactly the joint probability distribution over the set of variables:

$$P(X1, X2, \cdots, Xn) = \prod_{i=1}^{n} P(Xi \mid Pa(Xi))$$

This decomposition of a global function as a local product terms depending only of the node and its parents in the graph, is a fundamental property of Bayesian networks. It is the basis of the first work on the development of inference algorithms which calculate the probability of any variable of the model from the same partial observation of other variables. This has been proven NP-complete, but resulted in different algorithms that can be treated as information propagation methods in a graph. These methods obviously use the concept of conditional probability, i.e. what is the probability of $Xi$ knowing that I have observed $Xj$, but also the Bayes theorem, that calculates, conversely, the probability of $Xj$ knowing $Xi$, when $P(Xi \mid Xj)$ is known.

### B. Construction of a Bayesian network

As we have seen in the definition, the complete specification of a Bayesian network requires specifying a share structure (directed acyclic graph that underlies) and other parameters (probability tables). To do this, two approaches are possible and can be combined: the collection of expertise and machine learning, which is one of the attractions of Bayesian networks.

In the case of collection of expertise, the definition of the network structure begins with the identification of possible nodes and the distinction between (unobservable) informational variables (inputs) or hypothetical. The existence of an arc can be analyzed in terms of influence of one variable on another, but its orientation is more difficult. Traditionally, an arc is directed from A to B if A is a cause of B, but we will see that this interpretation is not as simple in the case of the learner modeling. The parameters are in turn attached in an approximate manner by using frequentists or qualitative information.

Since Bayesian network is a probability distribution, we can use maximum likelihood as statistical learning parameters criterion. The result is as a Bayesian network whose structure is fixed and E which is a comprehensive basis of example, the maximum likelihood is achieved if the parameters of the Bayesian network are equal to the frequencies of the same features observed in E. statistical learning structure requires for its development test to determine whether or not the random variables are conditionally independent [12].

## V. Development of Bayesian Network from a Use Case Diagramme

### A. The choice of Bayesian networks

As we previously presented, the diagrams of use cases is a top view of system features, it allows us to present all user actions (learner in our case). These actions may require elements of uncertainty, this uncertainty will clearly present when poised to collaboration diagrams. Representing this uncertainty becomes very important when there are a large number of interdependent and potentially conflicting requirements that overwhelm the capacity of spontaneous human spirit.

Bayesian network models explicitly the uncertainty between the requirements represented by use case and collaboration diagrams elements. During the presentation of the functional evidence such as the importance of a particular learner, a quantitative assessment can be performed to the way we strongly believe the requirement is indicated. We therefore see the ability to transform the use case diagrams of Bayesian networks as a significant potential lead in the modeling of the learner.

We believe that Bayesian networks will provide a solution that will allow us to understand and measure a dynamic way all the actions of the learner in a learning situation. Networks obtained, we will give a capacity to monitor and represent at real time, all the actions of the learner, the rationale for these choices, and identification of each of the paths that will be followed during a learning situation.

### B. Bayesian network development’s rules

#### 1) Generalization relationship

A generalized use case diagram contains a common functionality that is available for all the specialized use cases. The transformation of the generalization relationship to the nodes of a Bayesian network is simple:

Consider "Figure.4" in the use case A is a generalization of the use case A1 and the use case A2, we represent the functional requirement A1 and A2 being a descendant of the functional requirement of A.

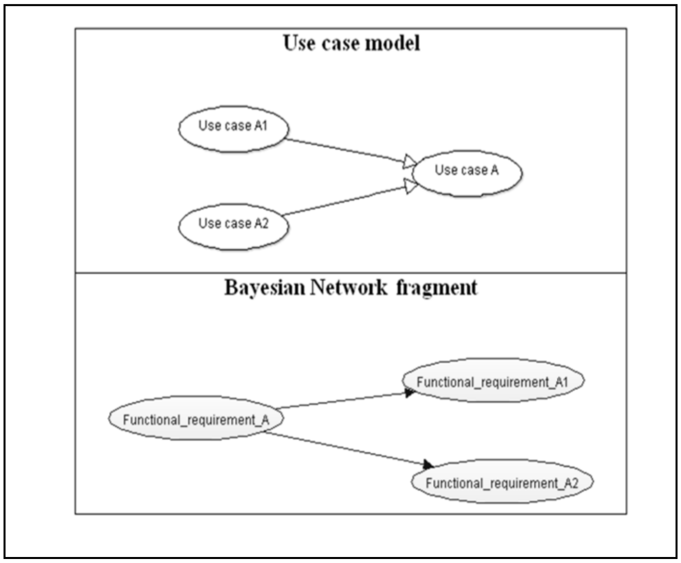
Use case model
Use case A1
Use case A
Use case A2
Bayesian Network fragment
Functional_requirement_A1
Functional_requirement_A
Functional_requirement_A2

Fig. 4. Bayesian fragment developed from the generalization relationship

This results in a Bayesian network with a similar structure. The direction of the arc flow from A to A1 and A2 reflecting a top-down decomposition. The information represented in the arrows of the use case will be included in the functional requirements. This indicates that it is more likely to encounter the general case of the specific functional requirement. Thus:

$$\mathrm{P(A) = prior}$$

$$\mathrm{P(A1 \mid A) \;=\; P(A \mid A1)P(A1) \;/\; P(A)}$$

$$\mathrm{P(A2 \mid A) \;=\; P(A \mid A21)P(A2) \;/\; P(A)}$$

*2) Inclusion relationship*

The inclusion relation in a use case diagram models the situation in which a use case is composed of a desired number of use cases. For inclusion, the high level of use cases cannot run without the implementation of sub use cases.

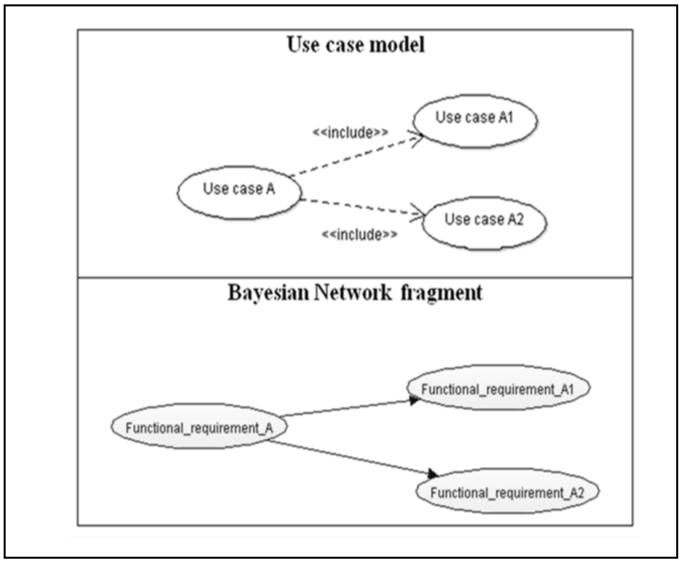

Fig. 5. Bayesian fragment developed from the inclusion relationship

To see how this can be translated to fragments of the Bayesian network, consider "Figure.5" use case A is linked to use cases A1 and A2 by an inclusion relation

This results in a Bayesian network with a similar structure as the generalization relationship. The direction of the arc flow from A to A1 and A2 reflecting a composition from bottom to top in the information represented in the arrows of the use case will be included in the functional requirements. This indicates that it is more likely to encounter the general case of the specific functional requirement. Thus:

$$P(A) = prior$$

$$P(\mathrm{A1 \mid A) \;=\; P(A \mid A1)P(A1) \;/\; P(A)}$$

$$P(\mathrm{A2 \mid A) \;=\; P(A \mid A21)P(A2) \;/\; P(A)}$$

*3) Extension relationship*

The extension relationship in a use case diagram represents a particular use case branched additional behavior given the satisfaction of certain conditions. In case of extension, the first use case does not need any more use case to run. The second use case is an exceptional behavior if the conditions are fulfilled.

Consider the general case schematized on "Figure 6". A use case is extended by the case of A1 use. This models the situation in which an additional criterion triggers the case of using A1 after executing use case A.

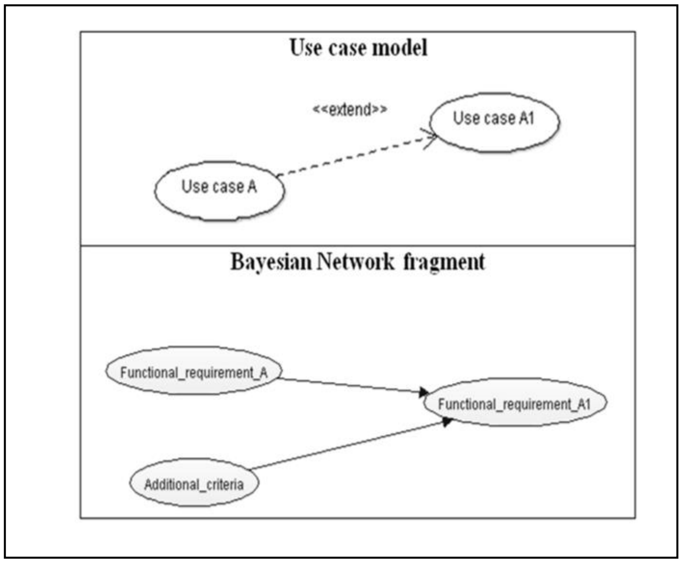

Fig. 6. Bayesian fragment developed from the extension relationship

The additional criterion is described in the flow of events from textual description. This situation is modeled as functional requirement A1 implied by the functional requirement A. The additional criterion is modeled as another functional requirement node. The direction of the implication is the additional criterion (AC) to the functional requirement A1. Thus:

$$\mathrm{P(A1 \mid A, AC) \;=\; \frac{P(A|A1, AC)P(A1|AC)}{P(A|AC)}}$$

## VI. Learner Modelling Using Bayesian Networks

Our work lies in the framework of learner modeling in an adaptive educational system, to illustrate the ideas discussed in the previous sections; we will focus our work on the actions of a learner in an adaptive system. We defined the "Table 1" several actions of a learner in a learning situation.

TABLE I. Learner Action in an Adaptive System

| Learner's actions |
|---|
| • Post question in the forum. |
| • Follow courses. |
| • Take the pretest. |

### A. Learner use case diagram

Considering "Figure.7" a main actor is identified, named the learner. The figure shows the generalization relationships between use cases and the learner, and generalization relationships, inclusion and extension between use cases.

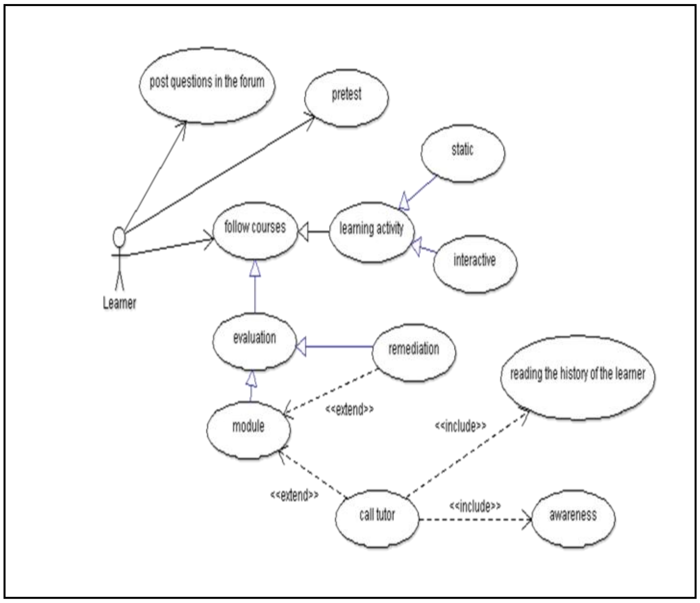

Fig. 7. Use case diagram of a learner model's actions in an adaptive system

In particular, the functional requirement "follow courses" is represented with a generalization relationship between the functional requirements "learning activity" and "evaluation". The functional requirements "post in the forum issues" and "pretest" are represented with a generalization relationship with actor "Learner". There are also extensions relationships in the functional requirement "module" and its relationship to functional requirements "remediation" and "call tutor." Inclusion relations are presented in the representation of the relationships between functional requirements "call tutor", "reading the history of the learner" and "system awareness."

### B. Bayesian network obtained

Once the use case diagrams have been created, it is easy to create the structure of the Bayesian network using the rules described in the previous sections. "Figure.8" represents the Bayesian network representation of the main actions of the learner in a Learning situation, constructed from the use case diagram shown in "Figure.7". Note how the conditional independence was directly modeled by applying the rules as shown.

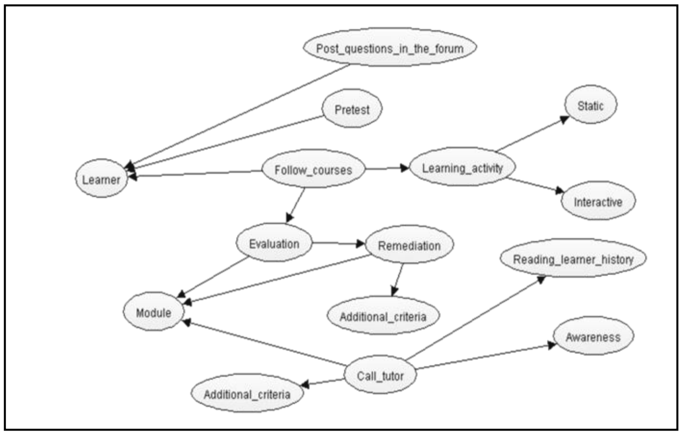

Fig. 8. Bayesian network developed from the learner model's use case diagram

## VII. Conclusion and Perspectives

We have shown in this work, how we can develop with well defined rules; Bayesian networks from use case diagrams of Unified Modeling Language, the development of the Bayesian network in future work could be done manually using the rules we provides in previous section of our work, or in using a software that can provide this transformation automatically, we couldn't find during our research a software which grantee this transformation, but we are working in development of a tools which allow this transformation automatically.

This work is a major step in our research in modeling the learner in an adaptive educational system, the transition from the use case diagrams towards Bayesian networks; give us the opportunity to reach our goal to use the Bayesian networks as a formalism to manage uncertainty in the modeling of the learner.

We see two main directions in which we can continue this work. On the one hand by applying our ideas to more advanced conceptual models as used so far. And on the other hand is transforming the Bayesian networks obtained a machine readable language, or one using probabilistic ontology as we proposed in previous work [13], or perform a combination of Bayesian networks with ontologies.

## References

[1] Booch, G., Rumbaugh, J. and Jacobson, I. The Unified Modeling Language User Guide. Addison-Wesley Publishing Company, Reading, MA, 1999.

[2] M. Anouar Tadlaoui, M. Khaldi, S. Aammou (2014) Towards a Learning model based on Bayesian Networks,EDULEARN14 Proceedings, pp. 3185-3193.

[3] M. Anouar Tadlaoui, M. Khaldi, S. Aammou (2014) Bayesian Networks for Learner Modeling, International Journal of Basic Sciences and Applied Computing 1 (1), 5-9.

[4] James Rumbaugh, Michael Blaha, William Premerlani, Frederick Eddy, William Lorensen (1990). Object-Oriented Modeling and Design. Prentice Hall

[5] Jacobson, Ivar; Grady Booch; James Rumbaugh (1998). The Unified Software Development Process. Addison Wesley Longman.

[6] L.Zaitseva, C.Boule. Learning systems in professional training. Workshop “Industry meets research” within the conference Interactive Computer Aided Learning ICL 2005 Villach, Austria 28 – 30 September 2005

[7] Beck, J., Stern, M., & Haugsjaa, E. (1996). Applications of AI in education. Crossroads, 3(1), 11-15.

[8] Brian Carr, Ira P. Goldstein. Overlays: a theory of modelling for computer aided instruction. AI Memo 406. 1977.

[9] Wenger, E. (1987). Artificial Intelligence and Tutoring Systems. Morgan Kaufmann, Los Altos, CA 94022.

[10] [PEARL, J., Probabilistic Reasoning in Intelligent System, Morgan Kaufmann. 1988.

[11] Geiger D., Heckermann D. (1996). Knowledge representation and inference in similarity networks and Bayesian multinets, Artificial Intelligence, volume 82 (1–2), pp. 45–74.

[12] BECKER, A., NAÏM, P., Les réseaux Bayésiens, modèles graphiques de connaissances, Eyrolles, 1999.

[13] M. Anouar Tadlaoui, M. Khaldi, S. Aammou (2014) Towards Probabilistic Ontology based on Bayesian Networks, International Journal of Software and Web Sciences 1 (10), 102-106.